TOXICOLOGICAL SCIENCES **136(1)**, 4–18 2013
doi:10.1093/toxsci/kft178
Advance Access publication August 19, 2013

Incorporating New Technologies Into Toxicity Testing and Risk Assessment: Moving From 21st Century Vision to a Data-Driven Framework

Russell S. Thomas,*[,1] Martin A. Philbert,† Scott S. Auerbach,‡ Barbara A. Wetmore,* Michael J. Devito,‡ Ila Cote,§ J. Craig Rowlands,¶ Maurice P. Whelan,‖ Sean M. Hays,‖‖ Melvin E. Andersen,* M. E. (Bette) Meek,‖‖‖ Lawrence W. Reiter,# Jason C. Lambert,** Harvey J. Clewell III,* Martin L. Stephens,†† Q. Jay Zhao,** Scott C. Wesselkamper,** Lynn Flowers,§ Edward W. Carney,¶ Timothy P. Pastoor,‡‡ Dan D. Petersen,** Carole L. Yauk,§§ and Andy Nong§§

**The Hamner Institutes for Health Sciences, Research Triangle Park, North Carolina 27709; †Toxicology Program, School of Public Health, University of Michigan, Ann Arbor, Michigan 48019; ‡Division of the National Toxicology Program, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, North Carolina 27709; §National Center for Environmental Assessment, U.S. Environmental Protection Agency, Arlington, Virginia 22202; ¶Toxicology & Environmental Research and Consulting, The Dow Chemical Company, Midland, Michigan 48674; ‖Systems Toxicology Unit, Institute for Health and Consumer Protection, European Commission DG Joint Research Centre, Ispra, Italy; ‖‖Summit Toxicology, L.L.P., Allenspark, Colorado 80510; ‖‖‖McLaughlin Centre for Population Health Risk Assessment, University of Ottawa, Ottawa, Ontario K1N6N5, Canada; #U.S. Environmental Protection Agency (Retired), Raleigh, North Carolina 27606; **National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, Ohio 45268; ††Johns Hopkins University, Bloomberg School of Public Health, Baltimore, Maryland 21205; ‡‡Toxicology & Health Sciences Department, Syngenta Crop Protection, Greensboro, North Carolina 27419; and §§Environmental Health Science and Research Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario K1A 0K9, Canada*

[1]To whom correspondence should be addressed at The Hamner Institutes for Health Sciences, 6 Davis Drive, PO Box 12137, Research Triangle Park, NC 27709. Fax: (919) 558-1300. E-mail: rthomas@thehamner.org.

Received June 12, 2013; accepted August 12, 2013

Based on existing data and previous work, a series of studies is proposed as a basis toward a pragmatic early step in transforming toxicity testing. These studies were assembled into a data-driven framework that invokes successive tiers of testing with margin of exposure (MOE) as the primary metric. The first tier of the framework integrates data from high-throughput *in vitro* assays, *in vitro*-to-*in vivo* extrapolation (IVIVE) pharmacokinetic modeling, and exposure modeling. The *in vitro* assays are used to separate chemicals based on their relative selectivity in interacting with biological targets and identify the concentration at which these interactions occur. The IVIVE modeling converts *in vitro* concentrations into external dose for calculation of the point of departure (POD) and comparisons to human exposure estimates to yield a MOE. The second tier involves short-term *in vivo* studies, expanded pharmacokinetic evaluations, and refined human exposure estimates. The results from the second tier studies provide more accurate estimates of the POD and the MOE. The third tier contains the traditional animal studies currently used to assess chemical safety. In each tier, the POD for selective chemicals is based primarily on endpoints associated with a proposed mode of action, whereas the POD for nonselective chemicals is based on potential biological perturbation. Based on the MOE, a significant percentage of chemicals evaluated in the first 2 tiers could be eliminated from further testing. The framework provides a risk-based and animal-sparing approach to evaluate chemical safety, drawing broadly from previous experience but incorporating technological advances to increase efficiency.

***Key Words:* *in vitro* and altenatives; biotransformation and toxicokinetics; predictive toxicology; risk assessment; safety evaluation; exposure.**

Shortly after the turn of the century, there was increasing recognition and acceptance within government agencies that new approaches were needed to evaluate the safety of the relatively large number of chemicals in commerce and the environment (EPA, 2003; Kavlock *et al.*, 2005; Meek and Armstrong, 2007; NTP, 2004). Following this recognition, the release of the National Research Council's Report "Toxicity Testing in the 21st Century: A Vision and a Strategy" (NRC, 2007) initiated a broad-based movement in the toxicology community to reassess how toxicity testing and risk assessment are performed. Since the release of the report, multiple efforts in the United States and abroad have added to the momentum with the shared goal of transitioning toxicity testing and risk assessment from an outdated, inefficient, costly, and animal-centric process to one that is more efficient, economical, less animal intensive, and more relevant to human health by utilizing new technologies that provide a better understanding of the underlying biological system. However, the majority of these efforts have focused more on a vision of how things should be done rather than the development of a pragmatic path forward that can be iteratively refined as greater understanding is achieved.

Over the past 5 years, a series of studies has been conducted that together may contribute first steps toward executing the shared vision. This article considers the implications of these studies and develops a new, data-driven tiered toxicity testing framework with potentially broad, international application across multiple regulatory agencies. The framework evolved from tiered approaches developed previously to address regulatory mandates to prioritize and assess large numbers of substances (Meek and Armstrong, 2007; Meek *et al.*, 2011). The primary application of the proposed framework is for chemicals with little or no safety-related data. In the United States, the proposed framework could, for example, be applied to new and legacy manufactured chemicals regulated under the Toxic Substances Control Act (TSCA). Conceptually, it could also be applied to other nonpharmaceutical chemicals that have defined toxicity testing requirements, such as pesticides. In Europe, the proposed framework could be applied to identify substances of concern within Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) that may have adverse effects that are not indicated in the basic information requirements linked to production volume.

TIER 1

The first tier of the proposed framework consists of 5 components—(1) use of high-throughput *in vitro* assays to separate chemicals into selective and nonselective modes of action; (2) *in vitro* genotoxicity assays to identify potential genotoxic and nongenotoxic chemicals; (3) *in vitro*-to-*in vivo* extrapolation (IVIVE) pharmacokinetic modeling to convert *in vitro* assay concentrations to applied doses; (4) high-throughput exposure modeling to estimate human exposures to chemicals; and (5) calculation of a margin of exposure (MOE) (Fig. 1). The data collection and analysis associated with each of these components would ideally occur simultaneously for inclusion in a Tier 1 data package associated with each chemical, analogous to the chemical dossiers currently being assembled for

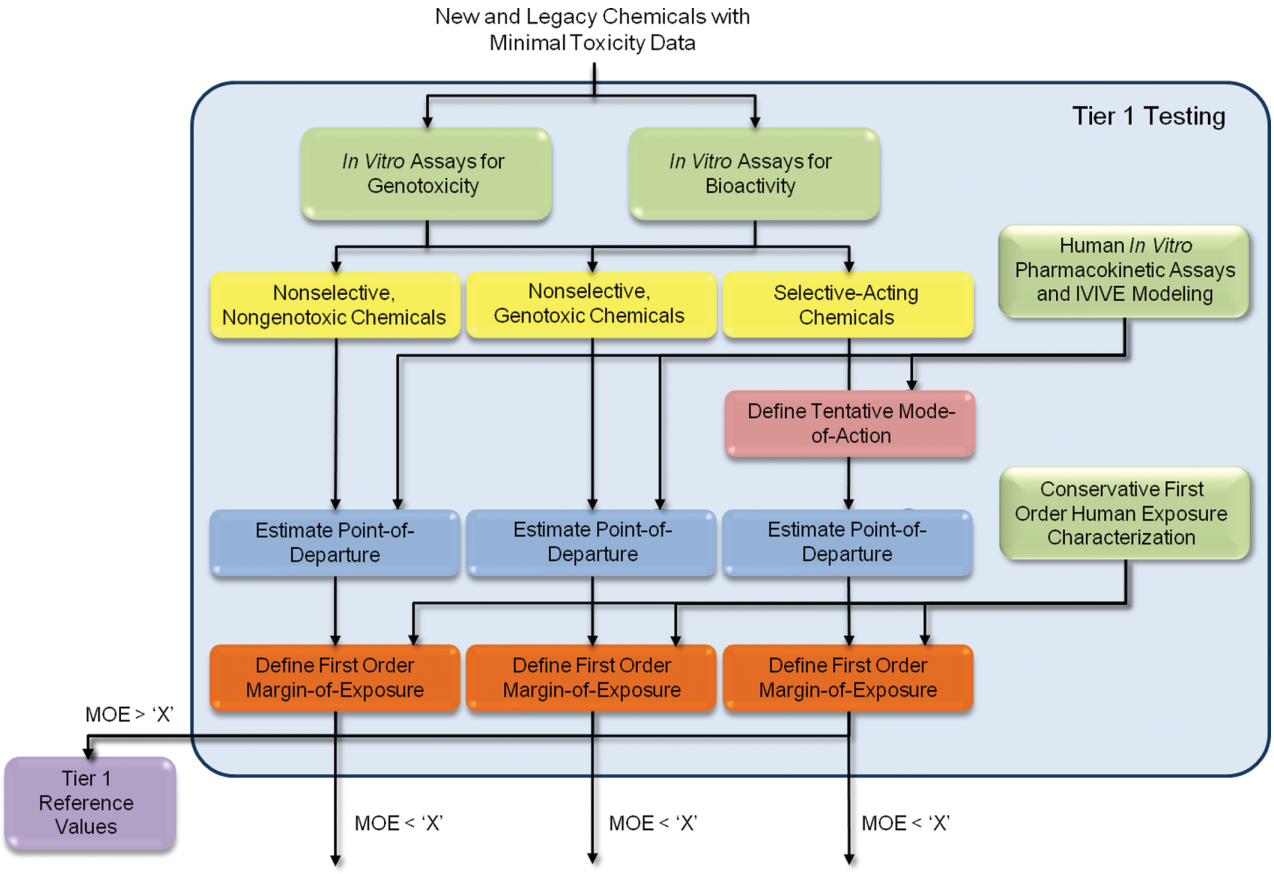

FIG. 1. A flowchart outlining Tier 1 in the proposed framework. The green boxes illustrate the Tier 1 data package that includes experimental data and computational modeling results that serve as inputs into the framework. The yellow boxes are separate chemical categories determined by the *in vitro* genotoxicity assays and the high-throughput *in vitro* screening assays. For the selective chemicals, the red box represents the determination of the tentative mode of action based on which high-throughput *in vitro* assays were selectively activated or inhibited. The blue and orange boxes represent the estimation of the point of departure and MOE using additional pharmacokinetic and exposure information, respectively. For those chemicals with a MOE greater than a defined cutoff, no further testing is performed, and Tier 1 reference values are published. Chemicals with a MOE less than the cutoff are advanced to Tier 2. Abbreviation: MOE, margin of exposure.

REACH. Within the current framework, the high-throughput *in vitro* screening would be composed of assays that are similar to those employed in the ToxCast project (Houck *et al.*, 2009; Judson *et al.*, 2010; Knudsen *et al.*, 2011; Martin *et al.*, 2010; Reif *et al.*, 2010; Rotroff *et al.*, 2010a). The *in vitro* ToxCast assays are primarily repurposed biochemical and cell-based assays used in drug discovery and cover 327 genes and 293 Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (Kavlock *et al.*, 2012). The details surrounding their selection and application have been discussed in the publications cited earlier. Although this selection of *in vitro* assays may not be ideal for covering the mechanistic landscape of chemically mediated toxicity, the ToxCast data set is the largest currently available for establishing the associations underlying this framework and provides a significant (though recognizably imperfect) starting point for a pragmatic path forward that can be refined as additional experience is acquired.

In Vitro *Assays for Bioactivity*

In the ToxCast phase I effort, 309 unique chemicals were screened in more than 600 *in vitro* assays. The 309 chemicals were primarily pesticides and high-production volume chemicals for which significant *in vivo* animal testing has been performed (Judson *et al.*, 2009; Martin *et al.*, 2009a). Although prediction of specific *in vivo* hazards with the current set of ToxCast *in vitro* assays is limited (Thomas *et al.*, 2012a; Wetmore *et al.*, 2012a), the *in vitro* assays do provide a broad biological profile of the potential proximal biochemical and cellular targets for a chemical. Instead of using the high-throughput *in vitro* assays to predict hazard, the initial role of high-throughput *in vitro* toxicity screening in our framework is to separate chemicals into either those that cause toxicity primarily through nonselective interactions with cells and cellular macromolecules or those that act through more selective interactions (eg, receptor-mediated chemicals).

To demonstrate how chemicals could be separated into these categories, the ToxCast phase I data were analyzed for surrogate measures of selectivity. Additional work in this area will need to be performed to identify which selective interactions are important for toxicity. The ToxCast *in vitro* assays were first grouped by gene to eliminate redundancy. For example, all of the *in vitro* assays evaluating the binding or transcriptional activation of estrogen receptor alpha (ESR1) were grouped together. The *in vitro* assays that did not correspond to a specific gene (eg, general cytotoxicity assays) were removed from the analysis. If any of the *in vitro* assays associated with a gene possessed an AC_{50} value (concentration at 50% of maximum activity), then the chemical was considered to have activity (ie, either activates or inhibits) toward that gene. Thus, the more genes with AC_{50} values, the less selective a chemical would be. For the ToxCast phase I chemicals, approximately 80% of chemicals have activity toward 10 or more genes, and the average chemical has activity toward approximately 20 genes (Fig. 2A).

A complementary measure of selectivity expands on this concept and takes into consideration the dose range across which the different *in vitro* assays were activated. For each chemical, the *in vitro* assays were again grouped by gene, and the *in vitro* assays that did not correspond to a specific gene were removed from the analysis (eg, general cytotoxicity assays). The potency of each chemical toward a specific gene was summarized by taking the minimum AC_{50} value among the associated *in vitro* assays. For example, if there were 5 *in vitro* assays evaluating the binding or transcriptional activation of ESR1, the potency

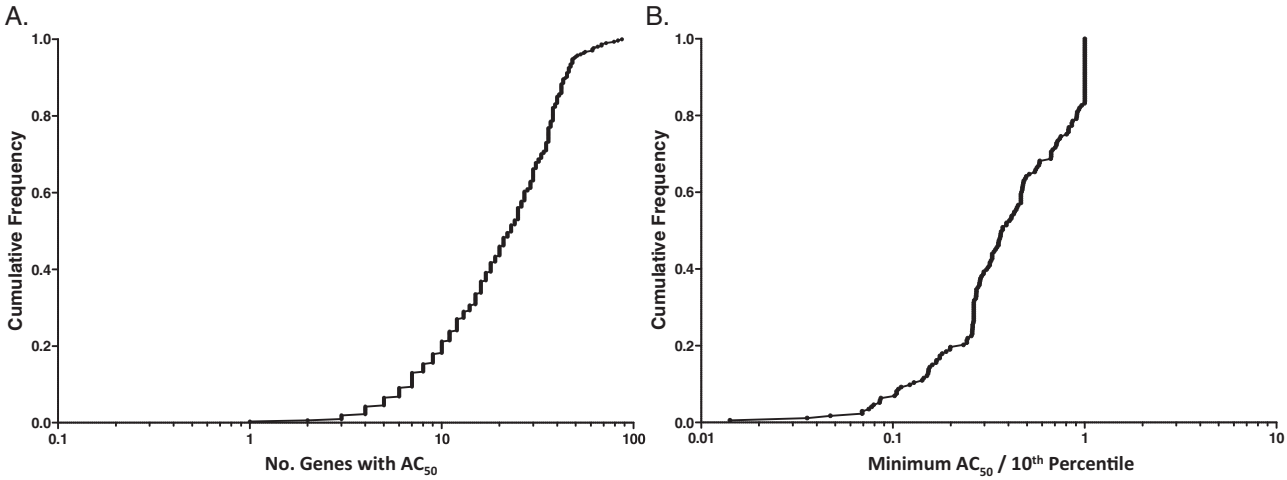

FIG. 2. Cumulative frequency plots summarizing the selectivity of the ToxCast phase I chemicals across the high-throughput *in vitro* assays. The *in vitro* assays were grouped by gene to eliminate redundancy. A, The number of genes with an AC_{50} value for each chemical. This represents the number of genes activated or inhibited by a chemical at any concentration. All 309 ToxCast phase I chemicals were used in this analysis. B, The ratio between the minimum AC_{50} and the 10th percentile for each chemical. This represents the ratio between the concentration at which the most sensitive gene-based assay is activated, and the concentration at which 10% of the gene-based assays are activated. Only 173 ToxCast phase I chemicals activating or inhibiting > 20 gene-based high-throughput *in vitro* assays were used in this analysis.

value for the most sensitive assay was selected. For each chemical, the summary values for each gene were then sorted, and the lowest summary value across all genes was divided by the 10th percentile. In other words, for each chemical, the potency of the *in vitro* assay for the most sensitive gene was divided by the relative potency when 10% of the genes were affected. According to this measure of selectivity, approximately 80% of the ToxCast phase I chemicals have a ratio of 3 or less between the most sensitive gene target and the 10th percentile (Fig. 2B). Taken together, these data suggest that the majority of chemicals represented in the ToxCast phase I library likely act via nonselective interactions with cellular macromolecules. These results may not be that surprising given that most chemicals in the ToxCast phase I library were nonpharmaceuticals and were developed either for their functional properties in a range of different products or pesticidal qualities and not optimized to interact specifically with a target protein.

In Vitro *Assays for Genotoxicity*

In addition to separating chemicals based on selectivity, chemicals are also grouped based on their genotoxic potential. This second screen was included as a practical consideration because most regulatory organizations either label or deal with substances that are likely to be carcinogenic through a mutagenic mode of action (MOA) differently than those considered not to be genotoxic. The *in vitro* assays for genotoxicity within the proposed framework include the bacterial reverse mutation test and the mammalian cell micronucleus test. These 2 *in vitro* genotoxicity assays cover the 3 critical genetic endpoints implicated in carcinogenesis and heritable diseases, namely gene mutations, and structural and numerical chromosome alterations. An expert panel convened by the European Food Safety Authority (EFSA) concluded that the inclusion of additional *in vitro* mammalian cell tests apart from these 2 assays would "significantly reduce specificity with no substantial gain in sensitivity" (EFSA, 2011). This is consistent with previous recommendations of multiple regulatory agencies and the International Conference on Harmonization (ICH) for registration of pharmaceuticals (HC, 2000). Adaptations have been proposed that reduce the number of false positives in the mammalian cell genotoxicity assays (Fowler *et al.*, 2012). If the 2 *in vitro* genotoxicity assays are negative, the chemical is placed in the nongenotoxic category, whereas positive results in both assays would classify the chemical as potentially genotoxic. If a chemical is positive in one of the *in vitro* assays, the chemical would be classified as potentially genotoxic and additional *in vitro* or *in vivo* assays could be performed to confirm the classification. For example, if a chemical is negative in the bacterial reverse mutation test and positive in the mammalian cell micronucleus test, an *in vivo* micronucleus analysis could also be performed to confirm the clastogenicity as recommended in the standard ICH guidelines (Müller *et al.*, 1999). Ideally, the *in vivo* micronuclei studies would be performed in Tier 2 on the same animals as the transcriptomic analysis using red blood cell micronuclei as the endpoint. If a chemical is positive in the bacterial reverse mutation test and negative in the mammalian cell micronucleus test, an *in vivo* gene mutation study could be performed using 28-day exposures of transgenic rodents carrying mutation reporter genes (OECD, 2011), but this is not cost effective for many chemicals at this point in time.

In Vitro *Pharmacokinetic Assays and IVIVE Modeling*

The internal dose of a chemical is an important determinant of toxicity. To incorporate pharmacokinetics into the framework, data from *in vitro* hepatic metabolic clearance and plasma protein binding assays are included in the Tier 1 data package. These 2 pharmacokinetic parameters are critical for estimating steady-state blood concentrations. In previous studies, experimental data from the *in vitro* pharmacokinetic assays were used to parameterize an IVIVE model to estimate the daily human oral dose, called the oral equivalent dose, necessary to produce steady-state *in vivo* blood concentrations equivalent to the *in vitro* AC_{50} value for each of the high-throughput *in vitro* assays (Rotroff *et al.*, 2010b; Wetmore *et al.*, 2012b). Oral equivalent dose values calculated using this approach are inherently conservative due to various assumptions in the IVIVE modeling (Rotroff *et al.*, 2010b; Wetmore *et al.*, 2012b). Monte Carlo sampling was also incorporated into the IVIVE model to account for interindividual variability and as a basis to derive an oral equivalent dose that represents the 95th percentile of the population. A similar approach takes place in our framework and when appropriate, other biokinetic factors within the *in vitro* assays will be taken into account (Blaauboer, 2010). Human equivalent concentrations could also be calculated to allow comparisons to air exposure levels.

Selective-Acting Chemicals and MOA

For the selective chemicals, the high-throughput *in vitro* assays are used to identify key events in a potential MOA. As previously defined, the MOA is a biologically plausible series of key events leading to an outcome that could be adverse (Sonich-Mullin *et al.*, 2001). Key events are empirically observable steps that are critical to the outcome. Although originally and often simply conceptualized as a linear series of key events, a MOA actually involves interdependent networks of events with feedback loops. A weight of evidence for the hypothesized MOA is established based on modified Bradford Hill considerations including dose response and temporal concordance of key events, consistency and specificity, and biological plausibility (Boobis *et al.*, 2006, 2008).

For application to selective chemicals in Tier 1 of the proposed framework, the key event must be triggered at doses lower than or equal to doses at which the adverse outcome is observed (ie, dose concordance between a key event and the adverse outcome). An example of this approach has been provided for the relationship between *in vitro* peroxisome proliferator–activated receptor alpha (PPARA) activation and rat liver proliferative lesions and liver tumors (Wetmore *et al.*, 2012a).

Among a subset of ToxCast phase I chemicals, 8 were active in the *in vitro* PPARA assays and cause proliferative lesions in the livers of rats, while 4 also cause rat liver tumors. For each chemical, the *in vivo*-derived lowest effect level (LEL) and no effect level for rat liver proliferative lesions and tumors were overlaid with the oral equivalent dose values for the 3 high-throughput *in vitro* assays measuring PPARA-related activation (Fig. 3). In all cases, activation of the PPARA *in vitro* assays occurred at or below the dose for the *in vivo* effects, thereby demonstrating dose concordance between this potential key event and the adverse outcome. A similar approach is taken in the proposed framework. From these selectively activated or inhibited assays, probable key event(s) will be determined and the oral equivalent dose value from a relevant *in vitro* assay will be compared with doses causing the associated *in vivo* adverse outcome (when available).

Apart from dose concordance, the biological relevance linking the *in vitro* assay to the *in vivo* adverse outcome will also be qualitatively assessed. If the *in vitro* assay representing the key event demonstrates dose concordance and biological relevance, the oral equivalent dose value will be used as the point of departure (POD) in a dose-response assessment. It is anticipated that there will be a limited number of MOAs that will be identified in Tier 1 for the selective chemicals. The MOAs will have a defined set of criteria based on chemicals that are known to act through those key events. Additional MOAs will be added as our biological knowledge improves and the assay suite is refined.

Nonselective Chemicals and Biological Perturbation

Identifying the MOA will generally be neither efficient nor economical for the nonselective chemicals because the chemicals will interact with and perturb multiple cellular processes. In a previous study, rat-specific oral equivalent doses were calculated for 59 ToxCast phase I chemicals and compared with LEL values for a variety of *in vivo* apical responses (Wetmore *et al.*, 2012a). Although not significantly correlated (r^2 = .046), the oral equivalent dose for the most sensitive *in vitro* assay was less than the LEL for the most sensitive rat *in vivo* endpoint for approximately 95% of the chemicals (Fig. 4A). The LEL values spanned 38 unique *in vivo* endpoints across multiple tissue and organ systems and across reproductive, chronic, and developmental study types. The median difference between the most sensitive *in vitro* assay and the most sensitive LEL value was 66-fold (Fig. 4B). These results indicate that the most sensitive *in vitro* assay was protective regardless of the nature of the adverse outcome and that bioactivity based on the oral equivalent dose for the most sensitive *in vitro* assay could be used as a reasonable estimate of the POD for the nonselective chemicals in a dose-response assessment.

Human Exposure Characterization

The final component of the Tier 1 data package relates to exposure. An understanding of human exposure is critical for placing the dose-response data for toxicity into context and ultimately assessing health risks. Human exposure estimates

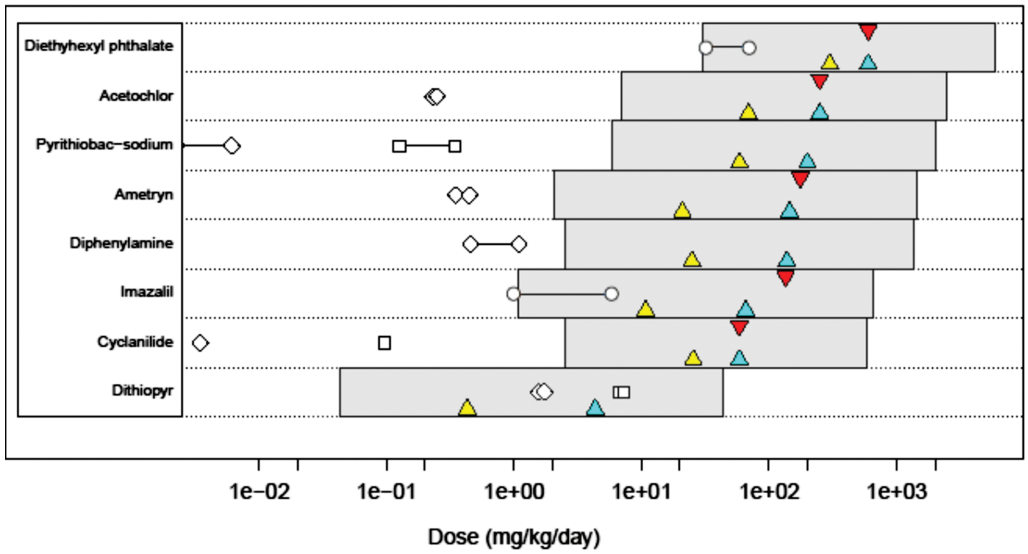

FIG. 3. Comparison of *in vivo* low effect and no effect doses for rat liver proliferative lesions and the *in vitro* PPARA-activating oral equivalent doses in the ToxCast phase I assays. The *in vivo* effects of 8 chemicals were plotted as LEL (blue triangles) and NEL (yellow triangles). If liver tumors were observed, a red triangle was placed at the tumorigenic LEL. The *in vitro* oral equivalent doses were overlaid for the PPARA transactivation assay in HepG2 cells (white circles) (Martin *et al.*, 2010) and PPARA-regulated HMGCS2 gene expression changes in primary human hepatocytes at 24 h (white squares) and 48 h (white diamonds) (Rotroff *et al.*, 2010a). The oral equivalent doses for each assay were calculated at the 1 and 10μM concentrations where intrinsic clearance was measured and connected with a solid line. The gray bars span 10-fold below the NEL and 10-fold above the highest of the LEL to illustrate potential uncertainties in the approximations associated with the oral equivalent doses. The graph and associated data are from Wetmore *et al.* (2012a). Abbreviations: LEL, low effect levels; NEL, no effect levels; PPARA, peroxisome proliferator–activated receptor alpha.

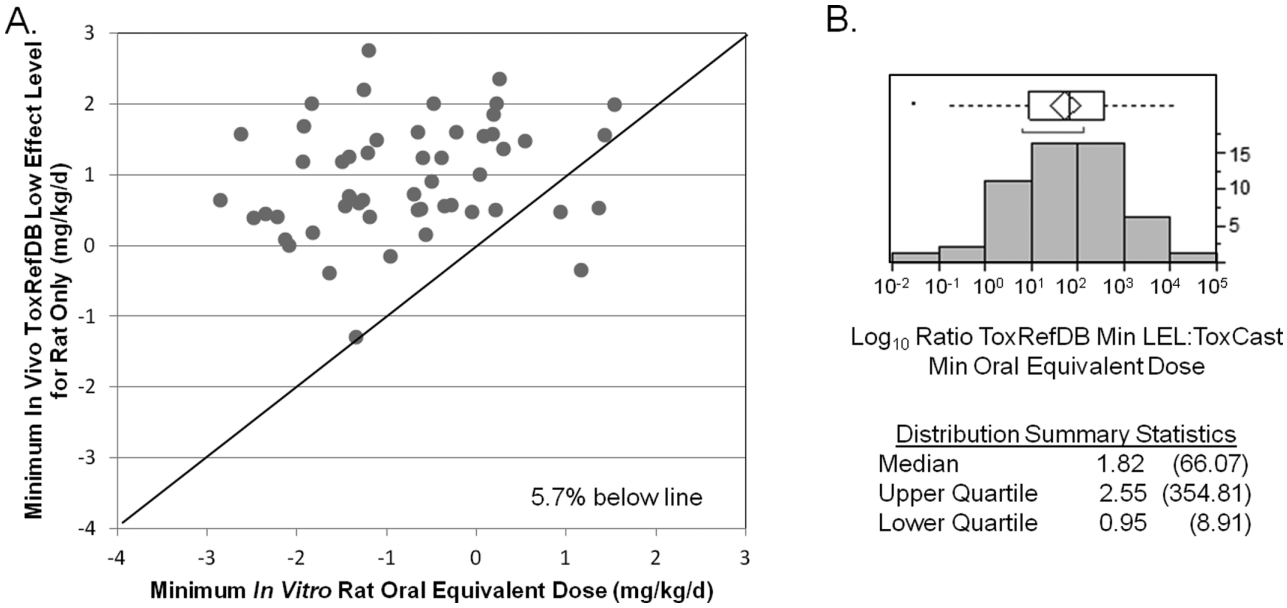

FIG. 4. A comparison of ToxCast *in vitro* assay possessing the lowest oral equivalent dose with the *in vivo* response possessing the lowest LEL for each chemical. A, Scatter plot of the log-transformed minimum LEL in ToxRefDB (y-axis) versus the minimum oral equivalent dose among the approximately 600 high-throughput *in vitro* ToxCast assays (x-axis) for each chemical. A subset of 59 ToxCast phase I chemicals were evaluated. B, Histogram and summary statistics of the $\log_{10}$-transformed ratios of the minimum LEL from the guideline animal studies in ToxRefDB divided by the minimum oral equivalent dose among the approximately 600 high-throughput *in vitro* assays. Nontransformed values for each of the summary statistics are provided in parentheses. The graph and associated data are from Wetmore *et al.* (2012a). Abbreviation: LEL, low effect level.

can be obtained from multiple sources and are associated with varying levels of uncertainty, depending principally on the complexity and accuracy of input data and model assumptions. Simple surrogates such as use and emission profiling and physicochemical properties offer potential to reasonably discriminate amongst chemicals with respect to exposure potential (Arnot *et al.*, 2012; Meek and Armstrong, 2007; Meek *et al.*, 2011).

For some of the Tier 1 chemicals, human exposure estimates can be derived from biomonitoring data such as the U.S. National Health and Nutrition Examination Survey (NHANES) (CDC, 2009). Measurements of parent chemical concentrations in the blood or urinary metabolite concentrations can be used to infer exposure concentrations using reverse dosimetry methods (Tan *et al.*, 2007). Using Monte Carlo sampling, probabilistic information about pharmacokinetics and exposure patterns can be included to estimate variability in the exposure distribution across a diverse population (Tan *et al.*, 2007). In Tier 1 of our framework, a conservative 95th percentile upper confidence bound on the exposure estimate is used, where available.

For chemicals without biomonitoring data, it is proposed to use a high-throughput exposure modeling that has been developed by combining existing environmental fate and transport models (ie, far field) with an empirical adjustment for indoor or consumer exposure (ie, near field) (Wambaugh *et al.*, 2013). In addition to a high-level assessment of consumer use, physicochemical properties and an estimate of environmental release are required as inputs to the model. The approach was calibrated using 82 chemicals from NHANES and then used to predict human exposures for 1,936 chemicals. The approach provides uncertainty bounds on the aggregate exposure estimates. The 95th percentile upper confidence bound of the exposure estimate is utilized in our framework.

MOEs

The human exposure estimates are combined with the POD from the dose-response assessment to calculate a MOE. For the components of the framework discussed above, the MOE is based on administered dose. A second alternative is the calculation of MOE based on internal dose using human biomonitoring data. For this alternative, concentrations in blood are directly compared with the concentrations in the *in vitro* assays so as to provide a direct measure of a MOE (Aylward and Hays, 2011; Aylward *et al.*, 2013). Although not all chemicals are amenable to biomonitoring measurements, the advantage of this approach is that it eliminates inherent uncertainty in the pharmacokinetic and exposure modeling. Interindividual variability could be incorporated by using either the percentiles of the measured distribution in large biomonitoring studies (CDC, 2009) or estimated from smaller studies using pooled samples (Caudill, 2010, 2012).

Regardless of whether the MOE is calculated using administered or internal dose, the MOE is the primary metric for determining whether a substance advances on for consideration in subsequent tiers. For the selective chemicals, the MOE is calculated relative to the oral equivalent dose from the *in vitro*

assay(s) associated with the proposed key event. For the nonselective chemicals, the MOE will be calculated relative to the oral equivalent dose for the most sensitive *in vitro* assay. For those chemicals with a MOE greater than some defined cutoff, no further testing would be required, whereas chemicals with a MOE below the cutoff would progress on to Tier 2. In a previous study on the ToxCast phase I chemicals (Wetmore *et al.*, 2012a), a Tier 1 MOE cutoff of > 100 would eliminate additional testing on approximately 40% of chemicals, whereas a MOE cutoff of > 1000 would eliminate additional testing on approximately 25% of chemicals (Fig. 5). Although the MOE cutoff will be primarily set through policy decisions and not necessarily based solely on scientific criteria, it should be set to provide adequate protection for susceptible subpopulations.

Reference Values

For those chemicals that do not advance to the next tier, Tier 1 reference values could be used to provide exposure guidelines. For certain regulatory agencies these could be considered screening-level values, whereas other agencies may wish to apply them in alternative ways. For the nongenotoxic chemicals, the POD and associated lower confidence bound can be divided by a set of uncertainty/safety factors to derive a Tier 1 reference value as a basis to provide health-related guidance. A previous study using *in vitro* data has provided one possible approach (Judson *et al.,* 2011). For the genotoxic chemicals, the POD and associated lower confidence bound can be divided by uncertainty/safety factors, or, alternatively, a linear, no threshold dose-response extrapolation could be applied. In this case, the POD could be assumed to be equivalent to a defined increase in adverse response (eg, 10%) and used to calculate a Tier 1 slope factor. However, it should be noted that activation of the *in vitro* assays analyzed here cannot yet be quantitatively linked to a specific incidence in the *in vivo* adverse response.

Chemicals Not Amenable to In Vitro *Screening*

Although efforts are underway to broaden the applicability of high-throughput *in vitro* screening platforms, certain chemicals, such as volatile substances, those not soluble in dimethyl sulfoxide or those that bind avidly to plastic, are currently not readily amenable for high-throughput *in vitro* screening. In the proposed framework, these chemicals would be identified using an initial structure-activity relationship or molecular properties filter and are automatically directed to Tier 2 for evaluation. No measure of selectivity would be available for these chemicals, and they would be assumed to act through nonselective mechanisms.

TIER 2

The second tier of the proposed framework consists of 5 components—(1) short-term *in vivo* transcriptomic studies to identify the transcriptional POD values for nonselective chemicals; (2) *in vivo* MOA studies to identify POD values for selective chemicals; (3) *in vitro* and *in vivo* pharmacokinetic studies to improve linkage between internal and applied dose; (4) refined human exposure estimates; and (5) calculation of a MOE (Fig. 6). Similar to the first tier, the data collection and

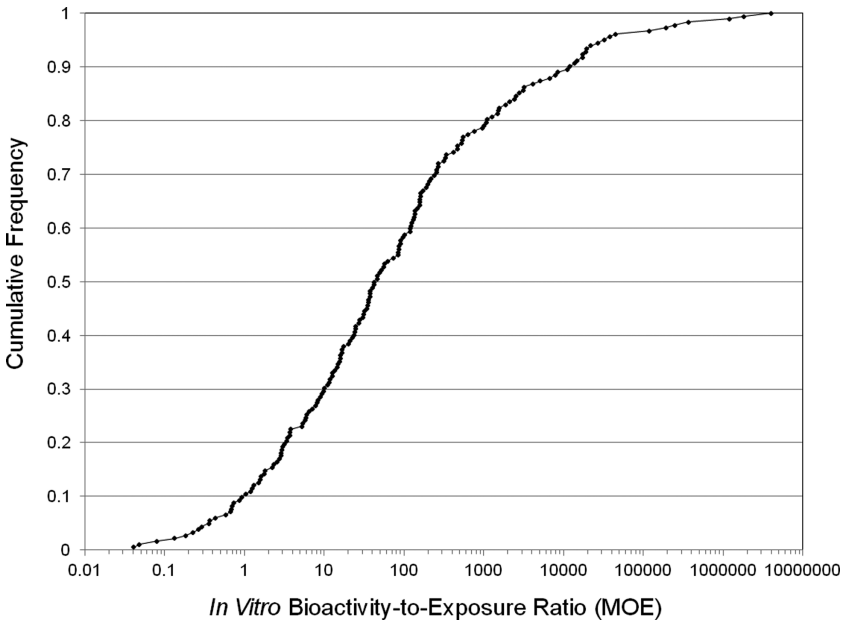

FIG. 5. Cumulative frequency plot of the ratio between *in vitro* bioactivity relative to human exposure in the most highly exposed subpopulation for the ToxCast phase I chemicals. The MOE was calculated by comparing the minimum oral equivalent dose across the approximately 600 high-throughput *in vitro* ToxCast assays using human pharmacokinetic measurements and *in vitro*-to-*in vivo* extrapolation modeling with the exposure estimates for the most highly exposed subpopulation. The graph and associated data are reproduced from Wetmore *et al.* (2012a). Abbreviation: MOE, margin of exposure.

analysis associated with each of these components would ideally occur simultaneously for inclusion in a cohesive Tier 2 data package. The data requirements would be different for the selective and nonselective chemicals.

Selective-Acting Chemicals and MOA

Focused *in vivo* MOA studies would be performed on the selective chemicals to evaluate the putative MOA defined in Tier 1 and would be defined by the type of MOA proposed. For proposed nuclear receptor agonists such as the constitutive androstane receptor, pregnane X receptor, or PPARA, these studies could include knockout and humanized rodent models. For proposed estrogen receptor agonists, these studies may also include standard *in vivo* studies such as the immature rat uterotrophic assay (OECD 440). Following the *in vivo* MOA studies and consideration of the weight of evidence, qualitative and quantitative human concordance would be evaluated (Boobis *et al.*, 2008). The POD for relevant MOAs would be determined based on the associated endpoints in the *in vivo* studies, but scaled, taking into account relevant kinetic and dynamic data.

Nonselective Chemicals and Biological Perturbation

The premise for analysis of nonselective chemicals has been developed based on the results of a series of previous studies examining the relationship between transcriptional and apical responses (Thomas *et al.*, 2007, 2011, 2012b, 2013). Both noncancer and cancer apical responses were analyzed using benchmark dose (BMD) methods to identify PODs. The dose response for changes in gene expression was also analyzed using BMD methods, and the responses were grouped based on signaling pathways. In a comparison of transcriptional BMD values for the most sensitive pathway with BMD values for the noncancer and cancer apical endpoints, there was a high degree of correlation (Thomas *et al.*, 2012b) (Fig. 7). The correlation was also robust across multiple time points (Thomas *et al.*, 2013). In this proposed framework, dose-response studies would be

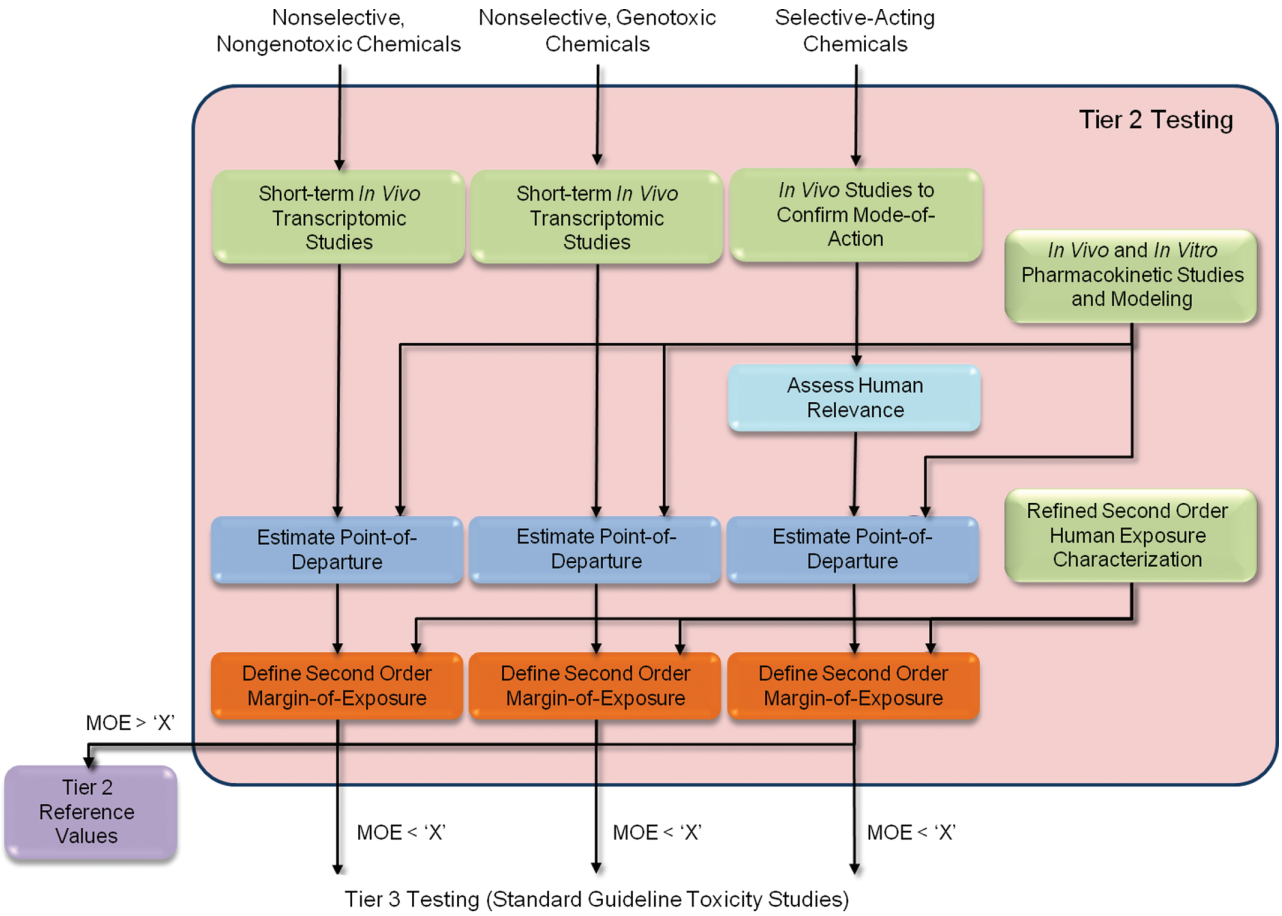

FIG. 6. A flowchart outlining Tier 2 in the proposed framework. The green boxes illustrate the Tier 2 data package that includes experimental data and computational modeling results that serve as inputs into the framework. The chemical categories determined by the *in vitro* genotoxicity assays and the high-throughput *in vitro* screening assays are retained from Tier 1. For the selective chemicals, the light blue box represents the determination of the human relevance of the mode of action. The blue and orange boxes represent the estimation of the point of departure and MOE using expanded pharmacokinetic and exposure information, respectively. For those chemicals with a MOE greater than a defined cutoff, no further testing is performed, and Tier 2 reference values are published. Chemicals with a MOE less than the cutoff are advanced to Tier 3. Abbreviation: MOE, margin of exposure.

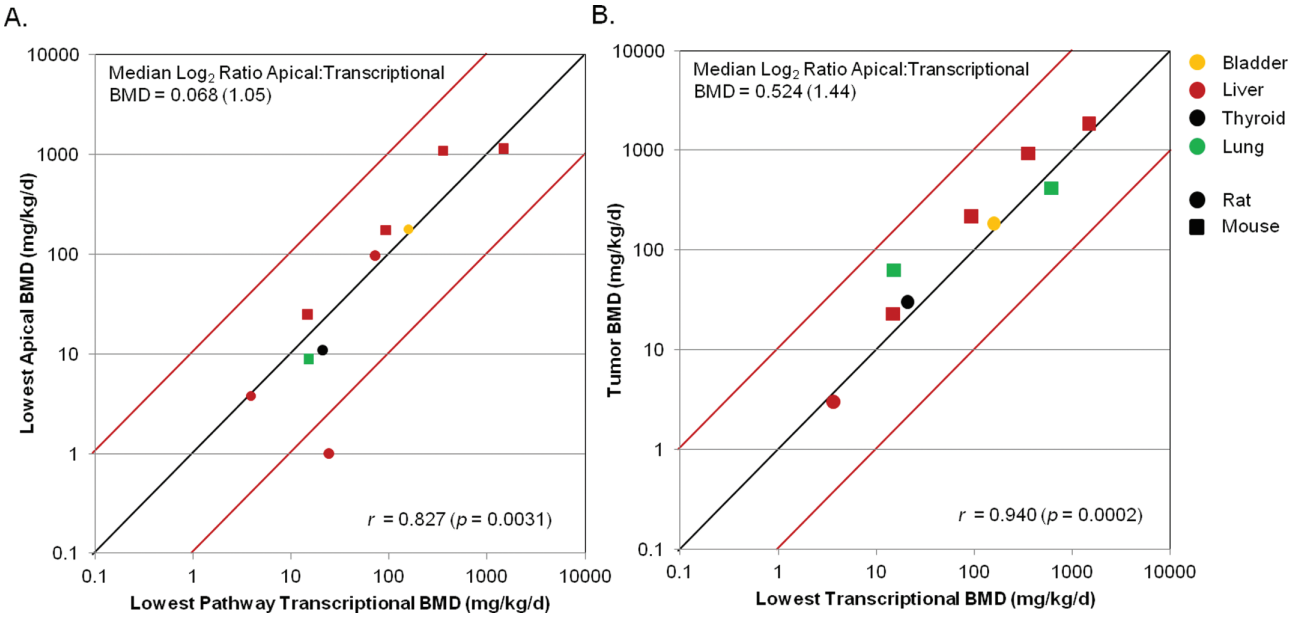

FIG. 7. Scatter plots of the relationship between the BMD values for (A) cancer-related apical endpoints or (B) noncancer apical endpoints and transcriptional BMD values for the most sensitive signaling pathway following 13 weeks of exposure. The data points were colored based on the target tissue, and symbol shape represents species. The black line dissecting the graph indicates equivalent apical and transcriptional BMD values. The red lines represent 10-fold difference between the BMD values. Correlation coefficients (r) and associated p value are included in the lower right hand corner of each panel. The median $\log_2$ ratio of the apical-to-transcriptional BMD values is provided in the upper left hand corner. The antilog of the transformed ratio is provided in parentheses. The graph and associated data are from Thomas *et al.* (2013). Abbreviation: BMD, benchmark dose.

performed on the chemical at any single time point between 5 days and 13 weeks in mice and rats of both sexes. A battery of 8 tissues that include those most frequently impacted in rodent cancer bioassays (liver, lung, mammary gland, stomach, vascular system, kidney, hematopoietic system, and urinary bladder) would be harvested. These 8 tissues cover 92% and 82% of targets for all mouse and rat carcinogens, respectively (Gold *et al.*, 2001). For noncancer effects, a previous analysis of 69 different pesticides demonstrated that adverse responses in a chronic rat bioassay generally occurred at lower doses than either a rat 2-generation reproductive study or a rat developmental toxicity study (Dourson *et al.*, 1992). In addition, potential reproductive and developmental toxicants could be identified based on consideration of weight of evidence of available data including structure activity modeling (Blackburn *et al.*, 2011). For chemicals flagged for these potential responses, additional tissues and alternate study designs (eg, OECD 421) would be incorporated to cover potential reproductive and developmental effects. Gene expression microarray analysis on these tissues would allow the estimation of pathway transcriptional BMD and BMDL values. The signaling pathway with the lowest transcriptional BMD value among the analyzed tissues would be used to derive the POD.

In Vivo *and* In Vitro *Pharmacokinetic Studies and Modeling*

The pharmacokinetic data from Tier 1 would also be expanded through focused *in vitro* studies and the collection of additional samples within the existing *in vivo* studies. Potential metabolites of each chemical would be identified using rodent and human microsomes, S9, and plasma. Recent advances in semiautomated metabolite identification have been made by combining *in silico* metabolite prediction software with ultra-high pressure liquid chromatography and mass spectral analysis (Bonn *et al.*, 2010). In essence, computational approaches are used to align the raw mass spectral data with a predicted site of metabolism in order to assign the structure of a metabolite. In our framework, these studies are performed to qualitatively identify probable metabolites in both the rodent and human. Reactive metabolites will also be evaluated using primary hepatocytes for analysis of lipid peroxidation, reactive oxygen species generation, and glutathione depletion.

Apart from the metabolite characterization, blood samples will be collected in either the *in vivo* mode-of-action studies for the selective chemicals or the *in vivo* transcriptomic studies for the nonselective chemicals. A previous study has demonstrated that the collection of blood samples at 3 specific time points during a single day of a repeat dose study was sufficient to estimate systemic bioavailability of a chemical (Saghir *et al.*, 2006). The human volume of distribution of each chemical will be estimated using quantitative structure-activity relationships (Peyret and Krishnan, 2011; Peyret *et al.*, 2010; Poulin and Theil, 2002; Rodgers and Rowland, 2006, 2007; Rodgers *et al.*, 2005), and the oral bioavailability will be estimated using the *in vitro* bidirectional permeability assay with Caco-2 cells (Wetmore *et al.*, 2012b). Together with the metabolic clearance and plasma protein binding data from the Tier 1 studies, the

combined data will be used to develop pharmacokinetic models to define the administered dose in humans leading to the active internal doses seen in the Tier 2 studies.

Human Exposure Characterization

Human exposure estimates from Tier 1 could potentially be refined through a combination of experimental measurements of physical-chemical properties and environmental degradation rates, refined release rates, detailed consumer and indoor use profiling, product formulation data, environmental and household exposure factor estimates, and targeted biomonitoring studies. The fate and transport models used for the high-throughput exposure modeling in Tier 1 rely heavily on physical-chemical properties, environmental half-lives, and release rates. For many chemicals, the physical-chemical properties and environmental half-lives are not experimentally available and have to be estimated using quantitative structure-activity relationships (Wambaugh *et al.*, 2013). Higher throughput methods are available for measuring both physical-chemical properties (Kerns, 2001) and the degradation half-lives in some environmental media (Hussain *et al.*, 2007). In our proposed framework, these higher throughput methods could be applied to all Tier 2 chemicals. In addition, more accurate industrial release rates could be obtained from industry. The combination of these data may be used to reduce the uncertainty in the fate and transport models.

To refine the near-field component of the exposure estimates, one option is to obtain more detailed consumer and indoor use profiles and product formulation data. These use profiles and formulation data could be aggregated from existing online databases such as the National Library of Medicine (NLM) Hazardous Substances Data Bank, NLM Household Products Database, Walmart's online material safety data sheets search engine, U.S. Environmental Protection Agency's (EPA) Exposure Factor Handbook (EPA, 2011), and the Substances in Products in the Nordic Countries (SPIN) database. A significant amount of the exposure-related data aggregation is currently underway (Egeghy *et al.*, 2012; Mattingly *et al.*, 2012). In the proposed framework, the more detailed use profile and product formulation data could potentially be used to select relevant consumer exposure models that exist for specific exposure scenarios (Kephalopoulos *et al.*, 2007). The resulting near-field exposure estimates could be combined with the far-field exposures to obtain an aggregated exposure estimate.

MOEs

Similar to Tier 1, the human exposure estimates will be combined with the POD from the *in vivo* dose-response studies to calculate a MOE based on administered dose. Within Tier 2, additional alternatives could be explored to integrate high-throughput biomonitoring studies in order to calculate a MOE based on internal dose. Advances in analytical chemistry and Fourier transform mass spectrometry have enabled what has been termed "high-performance metabolic profiling" (Johnson *et al.*, 2010; Park *et al.*, 2012; Soltow *et al.*, 2011). Using these methods, up to 7,000 chemicals can be measured in 20 µl of biofluids in 20 min (Soltow *et al.*, 2011). The high-performance metabolic profiling could be applied to biomonitoring samples from a range of different subpopulations, and mass labeled standards could be used to quantify specific chemicals of interest. The concentrations in the human blood samples would then be compared with estimated blood concentrations at the POD in the *in vivo* transcriptomic or MOA studies to provide a MOE.

Regardless of whether the MOE is calculated using administered or internal dose, the POD for the selective chemicals will be based on the associated endpoints in the *in vivo* MOA studies, whereas the POD for the nonselective chemicals will be derived from the signaling pathway with the lowest transcriptional BMD value across all analyzed tissues. For those chemicals with a MOE greater than a defined cutoff, no further testing would be required, whereas chemicals with a MOE below the cutoff would progress on to Tier 3. Using data from ToxRefDB on the *in vivo* LEL values for the ToxCast phase I chemicals (Knudsen *et al.*, 2009; Martin *et al.*, 2009a,b) and the corresponding human exposure estimates (Rotroff *et al.*, 2010b; Wetmore *et al.*, 2012b), MOE values were calculated using the minimum *in vivo* LEL value (ie, the most sensitive *in vivo* endpoint) and the exposure estimate for the most highly exposed subpopulation. Assuming that this ratio is representative of the MOE values that will be obtained from Tier 2 in our proposed framework, a MOE cutoff of > 100 would eliminate additional testing on approximately 97% of chemicals, whereas a MOE cutoff of > 1000 would eliminate testing on approximately 85% of chemicals (Fig. 8).

Reference Values

For those chemicals that are set aside from additional testing, Tier 2 reference values could be published to provide exposure guidelines. Similar to those described for Tier 1, the Tier 2 reference values may have different applications among regulatory agencies. For the nongenotoxic chemicals, the POD and associated lower confidence bound can be divided by a set of uncertainty/safety factors to derive a Tier 2 reference value as a basis to provide health-related guidance. For the genotoxic chemicals, the POD and associated lower confidence bound can be divided by uncertainty/safety factors, or, alternatively, a linear, no threshold dose-response extrapolation could be applied. The POD for the most sensitive pathway among the analyzed tissues could be used as the equivalent to the $BMDL_{10}$ for a tumor response (ie, the lower confidence limit of the BMD associated with a 10% increased risk of tumors). The rationale for this equivalence is that our data showed that transcriptional BMD values for the most sensitive pathway were, on average, within a factor of 2 of the BMD_{10} values for tumor responses in the corresponding rodent bioassays (Thomas *et al.*, 2013). The pathway transcriptional BMDL value could then be used to estimate a slope factor and the desired risk specific dose.

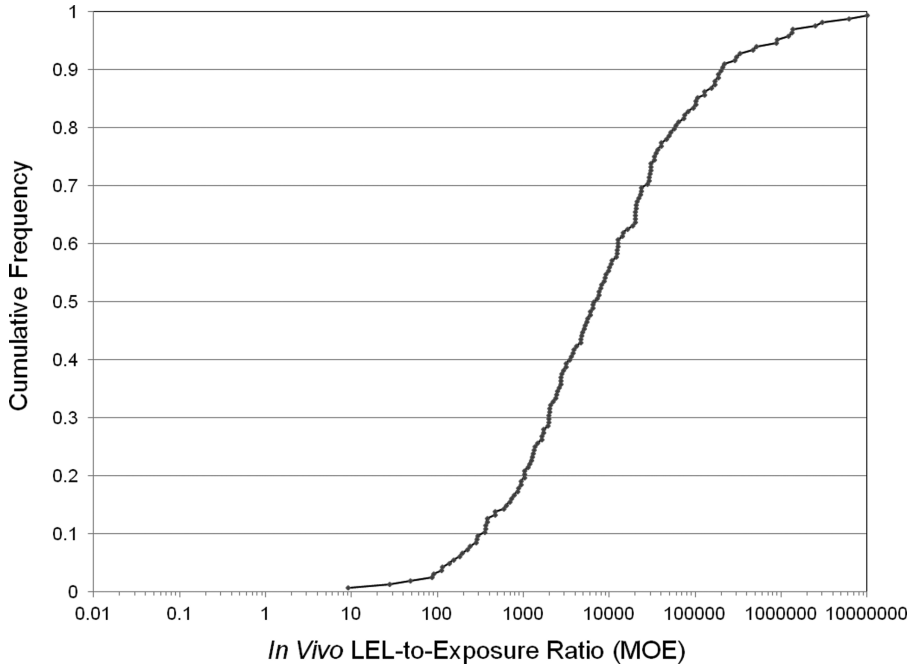

FIG. 8. Cumulative frequency plot of the ratio between the *in vivo* LEL relative to human exposure in the most highly exposed subpopulation for ToxCast phase I chemicals. The MOE was calculated by comparing the minimum LEL value of the guideline animal studies in ToxRefDB (Knudsen *et al.,* 2009; Martin *et al.*, 2009a,b) divided by the exposure estimates for the most highly exposed subpopulation (Wetmore *et al.*, 2012b). Abbreviations: LEL, low effect level; MOE, margin of exposure.

TIER 3

The toxicity testing proposed for chemicals passing into Tier 3 is not explicitly defined but would be conceptually equivalent to the traditional *in vivo* studies performed on high-value chemicals with significant potential for human exposure. Alternatively, they would be specified based on understanding of the toxicological profile acquired in lower tier testing. These studies could include rodent cancer bioassays, developmental toxicity studies, and reproductive toxicity studies. Depending on the MOE cutoff values imposed, it is anticipated that the majority of chemicals will be screened out in the preceding tiers and the estimated number of chemicals requiring these studies would be between 3% and 15%. Testing for these remaining compounds could be prioritized by endpoint and compounds based on the results from both Tier 1 and 2 studies.

SUMMARY

The proposed framework integrates data from new technologies into toxicity testing using what is currently the best available science. The proposed framework provides a risk-based, efficient, and animal-sparing approach to evaluate chemical safety. It is consistent with and draws broadly from previous experience in risk assessment but incorporates recent advances in technology to increase efficiency. The framework relies on the initial separation of chemicals into selective and nonselective MOAs. The prevailing thought is that the MOA approach can be efficiently applied to all chemicals. In our view, applying the MOA approach to the nonselective chemicals is neither practical, scientifically justifiable, nor consistent with evolving experience in increasing efficiency in risk assessment. First, imposing an MOA approach on these chemicals would waste valuable resources and unnecessarily delay decision making because each MOA requires agreement on the underlying key events followed by extensive peer review. Second, for chemicals which interact with numerous molecular targets in a cell, identifying a prevailing MOA among many operating in parallel will be a complex and potentially uncertain process. Thus, these efforts should be focused on the selective chemicals where a series of standardized MOAs could be developed that are associated with specific cellular targets.

A shift in thinking will also be required for the nonselective chemicals. The current hazard-based labeling approach that relies on apical responses will need to be transitioned to a "region of safety" approach where the most sensitive adverse apical effect is not known, but disruption of important biologic processes can be measured and the dose response characterized. The subsequent decisions are based on potential biological perturbations (Andersen and Krewski, 2010). This shift in thinking may be more challenging to implement due to the entrenched reliance on apical responses, the potentially conservative nature of utilizing molecular perturbations as PODs, and the difficulty in associating molecular perturbations with actual risk. To increase experience in using these approaches, the proposed framework could be tested in human health

risk assessment programs tasked with evaluating chemicals with limited toxicity data, such as the U.S. Environmental Protection Agency's (EPA) Superfund Health Risk Technical Support Center (STSC; http://www.epa.gov/superfund/health/research.htm). This program evaluates chemicals of interest to Superfund that commonly possess limited toxicity data compared with the generally more data-rich chemicals assessed by the EPA's Integrated Risk Information System (IRIS). Once a critical mass of chemicals has been evaluated, a reassessment of the framework would identify the strengths and limitations as well as incorporate any subsequent advances in technology or biological knowledge.

Several organizations have proposed using high-throughput *in vitro* assays for prioritizing chemicals for *in vivo* testing. In theory, under prioritization, all chemicals would still need to be tested, just in a different order. However, due to resource constraints, chemicals of lower concern may never be tested in practice. In the proposed framework, only a subset of chemicals would likely be elevated for additional testing in successive tiers based on the MOE, resulting in a significant monetary, time, and animal savings. However, it should be noted that depending on the needs of a given regulatory program, the appropriate "capture rate" for testing at each tier could be established, based on transparent delineation of considerations addressing this aspect as one of the components for acceptability of MOEs.

The proposed framework and underlying data sources also have some potential limitations that require improvement during the implementation of the framework. First, the current battery of *in vitro* assays may not represent one or more critical MOAs, which may falsely identify a chemical as nonselective. In this scenario, the POD for the chemical would be the most sensitive *in vitro* assay while the chemical may actually selectively activate or inhibit a pathway, process, or macromolecule at a significantly lower dose. In the proposed framework, it is important that the initial screen for selectivity be as comprehensive as possible. To broaden the biological space covered in Tier 1, additional *in vitro* assays could be included from those reported in PubChem (Li *et al.*, 2010). Alternatively, *in vitro* transcriptomic measurements could also be performed in concentration-response format across a panel of cell lines that express a diverse range of cellular targets and pathways (Lamb, 2007).

The second limitation that requires improvement is the properties associated with the current ToxCast *in vitro* assays. The limitations of these properties have been previously discussed (Judson *et al.*, 2011; Thomas *et al.*, 2012a) but include the lack of metabolism, use of single cell types that fail to replicate tissue-level cell-cell interactions, lack of biological context, and the short-term nature of the *in vitro* assays relative to responses following chronic duration exposure. As these assays are improved, the new assays can be incorporated into the Tier 1 screening process.

A third limitation is related to the exposure modeling across the tiers. In the current state of exposure modeling, the environmental fate and transport models (ie, far field) are generally more mature and have been more thoroughly validated than the indoor and consumer exposure models (ie, near field). Ironically, far-field exposures are usually the least important by orders of magnitude when considering the aggregate exposure to a particular chemical. The lack of well-validated indoor and consumer exposure models can be at least partially attributed to both a lack of data on the use and the relative quantity of specific chemicals in various consumer products and the lack of measurement data for model validation (Egeghy *et al.*, 2012). If exposure information is going to be used to make better decisions on chemical safety in either this framework or more generally, then the science relating to predicting indoor and consumer exposure needs to be improved.

A fourth limitation is adapting these new approaches to risk assessment of chemical mixtures. Current mixtures risk assessment methods focus either on component-based approaches or whole mixture toxicity evaluations (EPA, 2000). The framework described in this manuscript is readily adapted to a components-based approach. For example, for the selective chemicals that target the same protein or pathway, *in vivo* relative potency factors could be estimated using the results from the associated *in vitro* assays and IVIVE pharmacokinetic modeling to convert *in vitro* assay concentrations to applied doses. Mixtures risk estimates could then be derived using a relative potency factor approach similar to the toxic equivalency factors used for dioxins and dioxin-like chemicals (Van den Berg *et al.*, 2006). Alternatively, a hazard index approach could be applied to all chemicals of concern (EPA, 2000). A hazard quotient could be estimated by dividing the estimate of the human exposure by the Tier 1 reference value for each chemical. The cumulative hazard index is then estimated by summing the hazard quotients for all chemicals of concern. If the hazard index is > 1, this would indicate possible exposures of concern. Finally, the new *in vitro* approaches are also well suited for whole mixtures toxicity evaluation. The advantage of *in vitro* methods is that large numbers of relevant mixtures could be evaluated for biological activity. The proposed *in vitro* cumulative risk approach has many of the same limitations as the *in vitro* risk approach for individual chemicals. Although this proposed *in vitro* cumulative risk approach is theoretically possible, additional experimental data needs to be collected using this approach prior to application.

In other initiatives, frameworks have or are being developed that should be considered as complementary or congruent to that proposed here. For example, the Health and Environmental Sciences Institute (HESI) RISK21 project has proposed a more general and overarching scheme that easily incorporates the tiered evaluation approaches in this proposed framework. The WHO framework on MOA is currently being updated to reflect increasing experience in hypothesizing potential effects based on information pertaining to putative key events in established modes of action from appropriate *in vitro* or *in silico* systems and other evidence.

In summary, the proposed framework provides a risk-based, rapid, cost-effective, and animal-sparing means to evaluate chemicals for safety. The framework represents only an initial step that can be iteratively refined along the way to a fully knowledge-based approach for evaluating chemical safety. Progress toward a better future in chemical safety assessment will require implementation of pragmatic approaches, such as the proposed framework or one similar to it, by national regulatory agencies, cooperation of international organizations such as the Organization for Economic Co-operation and Development (OECD), and stakeholder education and involvement.

ACKNOWLEDGMENTS

The work by R.S.T., M.E.A., H.J.C. III, and B.A.W. on this manuscript was supported by the American Chemistry Council's Long-Range Research Initiative. The views expressed by authors affiliated with the U.S. Environmental Protection Agency are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency. This article may be the work product of an employee or group of employees of the National Institute of Environmental Health Sciences (NIEHS), National Institutes of Health (NIH); however, the statements, opinions, or conclusions contained therein do not necessarily represent the statements, opinions, or conclusions of NIEHS, NIH, or the U.S. government. The authors wish to thank Chad Blystone and Alex Merrick for reviewing the manuscript.

REFERENCES

Andersen, M. E., and Krewski, D. (2010). The vision of toxicity testing in the 21st century: Moving from discussion to action. *Toxicol. Sci.* **117,** 17–24.

Arnot, J. A., Brown, T. N., Wania, F., Breivik, K., and McLachlan, M. S. (2012). Prioritizing chemicals and data requirements for screening-level exposure and risk assessment. *Environ. Health Perspect*. **120,** 1565–1570.

Aylward, L. L., and Hays, S. M. (2011). Consideration of dosimetry in evaluation of ToxCast™ data. *J. Appl. Toxicol*. **31,** 741–751.

Aylward, L. L., Kirman, C. R., Schoeny, R., Portier, C. J., and Hays, S. M. (2013). Evaluation of biomonitoring data from the CDC National Exposure Report in a risk assessment context: Perspectives across chemicals. *Environ. Health Perspect*. **121,** 287–294.

Blaauboer, B. J. (2010). Biokinetic modeling and in vitro-in vivo exptrapolations. *J. Toxicol. Environ. Health. Part B* **13,** 242–252.

Blackburn, K., Bjerke, D., Daston, G., Felter, S., Mahony, C., Naciff, J., Robison, S., and Wu, S. (2011). Case studies to test: A framework for using structural, reactivity, metabolic and physicochemical similarity to evaluate the suitability of analogs for SAR-based toxicological assessments. *Regul. Toxicol. Pharmacol*. **60,** 120–135.

Bonn, B., Leandersson, C., Fontaine, F., and Zamora, I. (2010). Enhanced metabolite identification with MS(E) and a semi-automated software for structural elucidation. *Rapid Commun. Mass Spectrom*. **24,** 3127–3138.

Boobis, A. R., Cohen, S. M., Dellarco, V., McGregor, D., Meek, M. E., Vickers, C., Willcocks, D., and Farland, W. (2006). IPCS framework for analyzing the relevance of a cancer mode of action for humans. *Crit. Rev. Toxicol*. **36,** 781–792.

Boobis, A. R., Doe, J. E., Heinrich-Hirsch, B., Meek, M. E., Munn, S., Ruchirawat, M., Schlatter, J., Seed, J., and Vickers, C. (2008). IPCS framework for analyzing the relevance of a noncancer mode of action for humans. *Crit. Rev. Toxicol*. **38,** 87–96.

Caudill, S. P. (2010). Characterizing populations of individuals using pooled samples. *J. Expo. Sci. Environ. Epidemiol*. **20,** 29–37.

Caudill, S. P. (2012). Use of pooled samples from the National Health and Nutrition Examination Survey. *Stat. Med*. **31,** 3269–3277.

CDC (2009). Fourth National Report on Human Exposure to Environmental Chemicals (N. C. f. H. Statistics, Ed.) Centers for Disease Control and Prevention, Atlanta, GA.

Dourson, M. L., Knauf, L. A., and Swartout, J. C. (1992). On reference dose (RfD) and its underlying toxicity data base. *Toxicol. Ind. Health* **8,** 171–189.

EFSA (2011). Scientific opinion on genotoxicity testing strategies applicable to food and feed safety assessment. *EFSA J*. **9,** 2379–2447.

Egeghy, P. P., Judson, R., Gangwal, S., Mosher, S., Smith, D., Vail, J., and Cohen Hubal, E. A. (2012). The exposure data landscape for manufactured chemicals. *Sci. Total Environ*. **414,** 159–166.

EPA (2000). *Supplementary Guidance for Conducting Health Risk Assessment of Chemical Mixtures*. U.S. Environmental Protection Agency, Washington, DC. EPA/630/R-00/002.

EPA (2003). *A Framework for a Computational Toxicology Research Program in ORD*. U.S. Environmental Protection Agency, Washington, DC. EPA/600/R-03/065.

EPA (2011). *Exposure Factors Handbook*. U.S. Environmental Protection Agency, Washington, DC. EPA/600/R-09/052F.

Fowler, P., Smith, R., Smith, K., Young, J., Jeffrey, L., Kirkland, D., Pfuhler, S., and Carmichael, P. (2012). Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. II. Importance of accurate toxicity measurement. *Mutat. Res*. **747,** 104–117.

Gold, L. S., Manley, N. B., Slone, T. H., and Ward, J. M. (2001). Compendium of chemical carcinogens by target organ: Results of chronic bioassays in rats, mice, hamsters, dogs, and monkeys. *Toxicol. Pathol*. **29,** 639–652.

HC (2000). *Genotoxicity: A Standard Batter for Genotoxicity Testing of Pharmaceuticals*. Therapeutic Products Program, Ottawa, Canada. ICH Topic S2B.

Houck, K. A., Dix, D. J., Judson, R. S., Kavlock, R. J., Yang, J., and Berg, E. L. (2009). Profiling bioactivity of the ToxCast chemical library using BioMAP primary human cell systems. *J. Biomol. Screen*. **14,** 1054–1066.

Hussain, S., Arshahd, M., Saleem, M., and Zahir, Z. A. (2007). Screening of soil fungi for in vitro degradation of endosulfan. *World J. Microbiol. Biotechnol*. **23,** 939–945.

Johnson, J. M., Yu, T., Strobel, F. H., and Jones, D. P. (2010). A practical approach to detect unique metabolic patterns for personalized medicine. *Analyst* **135,** 2864–2870.

Judson, R., Richard, A., Dix, D. J., Houck, K., Martin, M., Kavlock, R., Dellarco, V., Henry, T., Holderman, T., Sayre, P., *et al.* (2009). The toxicity data landscape for environmental chemicals. *Environ. Health Perspect*. **117,** 685–695.

Judson, R. S., Houck, K. A., Kavlock, R. J., Knudsen, T. B., Martin, M. T., Mortensen, H. M., Reif, D. M., Rotroff, D. M., Shah, I., Richard, A. M., *et al.* (2010). In vitro screening of environmental chemicals for targeted testing prioritization: The ToxCast project. *Environ. Health Perspect*. **118,** 485–492.

Judson, R. S., Kavlock, R. J., Setzer, R. W., Hubal, E. A., Martin, M. T., Knudsen, T. B., Houck, K. A., Thomas, R. S., Wetmore, B. A., and Dix, D. J. (2011). Estimating toxicity-related biological pathway altering doses for high-throughput chemical risk assessment. *Chem. Res. Toxicol*. **24,** 451–462.

Kavlock, R., Ankley, G. T., Collette, T., Francis, E., Hammerstrom, K., Fowle, J., Tilson, H., Toth, G., Schmieder, P., Veith, G. D., *et al.* (2005). Computational toxicology: Framework, partnerships, and program development. September 29-30, 2003, Research Triangle Park, North Carolina. *Reprod. Toxicol*. **19,** 265–280.

Kavlock, R., Chandler, K., Houck, K., Hunter, S., Judson, R., Kleinstreuer, N., Knudsen, T., Martin, M., Padilla, S., Reif, D., *et al.* (2012). Update on EPA's ToxCast program: Providing high throughput decision support tools for chemical risk management. *Chem. Res. Toxicol*. **25,** 1287–1302.

Kephalopoulos, S., Bruinen de Bruin, Y., Arvanitis, A., Hakkinen, P., and Jantunen, M. (2007). Issues in consumer exposure modeling: Towards harmonization on a global scale. *J. Expo. Sci. Environ. Epidemiol*. **17**(Suppl. 1), S90–S100.

Kerns, E. H. (2001). High throughput physicochemical profiling for drug discovery. *J. Pharm. Sci*. **90,** 1838–1858.

Knudsen, T. B., Houck, K. A., Sipes, N. S., Singh, A. V., Judson, R. S., Martin, M. T., Weissman, A., Kleinstreuer, N. C., Mortensen, H. M., Reif, D. M., *et al.* (2011). Activity profiles of 309 ToxCast™ chemicals evaluated across 292 biochemical targets. *Toxicology* **282,** 1–15.

Knudsen, T. B., Martin, M. T., Kavlock, R. J., Judson, R. S., Dix, D. J., and Singh, A. V. (2009). Profiling the activity of environmental chemicals in prenatal developmental toxicity studies using the U.S. EPA's ToxRefDB. *Reprod. Toxicol*. **28,** 209–219.

Lamb, J. (2007). The Connectivity Map: A new tool for biomedical research. *Nat. Rev. Cancer* **7,** 54–60.

Li, Q., Cheng, T., Wang, Y., and Bryant, S. H. (2010). PubChem as a public resource for drug discovery. *Drug Discov. Today* **15,** 1052–1057.

Martin, M. T., Dix, D. J., Judson, R. S., Kavlock, R. J., Reif, D. M., Richard, A. M., Rotroff, D. M., Romanov, S., Medvedev, A., Poltoratskaya, N., *et al.* (2010). Impact of environmental chemicals on key transcription regulators and correlation to toxicity end points within EPA's ToxCast program. *Chem. Res. Toxicol*. **23,** 578–590.

Martin, M. T., Judson, R. S., Reif, D. M., Kavlock, R. J., and Dix, D. J. (2009a). Profiling chemicals based on chronic toxicity results from the U.S. EPA ToxRef Database. *Environ. Health Perspect*. **117,** 392–399.

Martin, M. T., Mendez, E., Corum, D. G., Judson, R. S., Kavlock, R. J., Rotroff, D. M., and Dix, D. J. (2009b). Profiling the reproductive toxicity of chemicals from multigeneration studies in the toxicity reference database. *Toxicol. Sci*. **110,** 181–190.

Mattingly, C. J., McKone, T. E., Callahan, M. A., Blake, J. A., and Hubal, E. A. (2012). Providing the missing link: The exposure science ontology ExO. *Environ. Sci. Technol*. **46,** 3046–3053.

Meek, M. E., and Armstrong, V. C. (2007). The management of industrial chemicals in Canada. In *Risk Assessment of Chemicals: An introduction* (C. J. Van Leeuwen and T. G. Vermeire, Eds.), pp. 591–622. Springer, Dordrecht, Netherlands.

Meek, M. E., Boobis, A. R., Crofton, K. M., Heinemeyer, G., Raaij, M. V., and Vickers, C. (2011). Risk assessment of combined exposure to multiple chemicals: A WHO/IPCS framework. *Regul. Toxicol. Pharmacol*. **60,** S1–S14.

Müller, L., Kikuchi, Y., Probst, G., Schechtman, L., Shimada, H., Sofuni, T., and Tweats, D. (1999). ICH-harmonised guidances on genotoxicity testing of pharmaceuticals: Evolution, reasoning and impact. *Mutat. Res*. **436,** 195–225.

NRC (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy*. National Research Council of the National Academies, Washington, DC.

NTP (2004). *A National Toxicology Program for the 21st Century: A Roadmap for the Future*. National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC.

OECD (2011). Test No. 488: Transgenic Rodent Somatic and Germ Cell Gene Mutation Assays. In *OECD Guidelines for the Testing of Chemicals, Section 4*. Organisation for Economic Co-operation and Development, Paris, France.

Park, Y. H., Lee, K., Soltow, Q. A., Strobel, F. H., Brigham, K. L., Parker, R. E., Wilson, M. E., Sutliff, R. L., Mansfield, K. G., Wachtman, L. M., *et al.* (2012). High-performance metabolic profiling of plasma from seven mammalian species for simultaneous environmental chemical surveillance and bioeffect monitoring. *Toxicology* **295,** 47–55.

Peyret, T., and Krishnan, K. (2011). QSARs for PBPK modelling of environmental contaminants. *SAR QSAR Environ. Res*. **22,** 129–169.

Peyret, T., Poulin, P., and Krishnan, K. (2010). A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. *Toxicol. Appl. Pharmacol*. **249,** 197–207.

Poulin, P., and Theil, F. P. (2002). Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. *J. Pharm. Sci*. **91,** 129–156.

Reif, D. M., Martin, M. T., Tan, S. W., Houck, K. A., Judson, R. S., Richard, A. M., Knudsen, T. B., Dix, D. J., and Kavlock, R. J. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environ. Health Perspect*. **118,** 1714–1720.

Rodgers, T., Leahy, D., and Rowland, M. (2005). Physiologically based pharmacokinetic modeling 1: Predicting the tissue distribution of moderate-to-strong bases. *J. Pharm. Sci*. **94,** 1259–1276.

Rodgers, T., and Rowland, M. (2006). Physiologically based pharmacokinetic modelling 2: Predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *J. Pharm. Sci*. **95,** 1238–1257.

Rodgers, T., and Rowland, M. (2007). Mechanistic approaches to volume of distribution predictions: Understanding the processes. *Pharm. Res*. **24,** 918–933.

Rotroff, D. M., Beam, A. L., Dix, D. J., Farmer, A., Freeman, K. M., Houck, K. A., Judson, R. S., LeCluyse, E. L., Martin, M. T., Reif, D. M., *et al.* (2010a). Xenobiotic-metabolizing enzyme and transporter gene expression in primary cultures of human hepatocytes modulated by ToxCast chemicals. *J. Toxicol. Environ. Health. B. Crit. Rev*. **13,** 329–346.

Rotroff, D. M., Wetmore, B. A., Dix, D. J., Ferguson, S. S., Clewell, H. J., Houck, K. A., Lecluyse, E. L., Andersen, M. E., Judson, R. S., Smith, C. M., *et al.* (2010b). Incorporating human dosimetry and exposure into high-throughput in vitro toxicity screening. *Toxicol. Sci*. **117,** 348–358.

Saghir, S. A., Mendrala, A. L., Bartels, M. J., Day, S. J., Hansen, S. C., Sushynski, J. M., and Bus, J. S. (2006). Strategies to assess systemic exposure of chemicals in subchronic/chronic diet and drinking water studies. *Toxicol. Appl. Pharmacol*. **211,** 245–260.

Soltow, Q. A., Strobel, F. H., Mansfield, K. G., Wachtman, L. M., Park, Y., and Jones, D. P. (2011). High-performance metabolic profiling with dual chromatography-Fourier-transform mass spectrometry (DC-FTMS) for study of the exposome. *Metabolomics* **9,** 132–143.

Sonich-Mullin, C., Fielder, R., Wiltse, J., Baetcke, K., Dempsey, J., Fenner-Crisp, P., Grant, D., Hartley, M., Knaap, A., Kroese, D., *et al.*; International Programme on Chemical Safety. (2001). IPCS conceptual framework for evaluating a mode of action for chemical carcinogenesis. *Regul. Toxicol. Pharmacol*. **34,** 146–152.

Tan, Y. M., Liao, K. H., and Clewell, H. J., 3rd. (2007). Reverse dosimetry: Interpreting trihalomethanes biomonitoring data using physiologically based pharmacokinetic modeling. *J. Expo. Sci. Environ. Epidemiol*. **17,** 591–603.

Thomas, R. S., Allen, B. C., Nong, A., Yang, L., Bermudez, E., Clewell, H. J., 3rd, and Andersen, M. E. (2007). A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. *Toxicol. Sci*. **98,** 240–248.

Thomas, R. S., Black, M. B., Li, L., Healy, E., Chu, T. M., Bao, W., Andersen, M. E., and Wolfinger, R. D. (2012a). A comprehensive statistical analysis of predicting in vivo hazard using high-throughput in vitro screening. *Toxicol. Sci*. **128,** 398–417.

Thomas, R. S., Clewell, H. J., 3rd, Allen, B. C., Wesselkamper, S. C., Wang, N. C., Lambert, J. C., Hess-Wilson, J. K., Zhao, Q. J., and Andersen, M. E. (2011). Application of transcriptional benchmark dose values in quantitative cancer and noncancer risk assessment. *Toxicol. Sci*. **120,** 194–205.

Thomas, R. S., Clewell, H. J., 3rd, Allen, B. C., Yang, L., Healy, E., and Andersen, M. E. (2012b). Integrating pathway-based transcriptomic data into quantitative chemical risk assessment: A five chemical case study. *Mutat. Res*. **746,** 135–143.

Thomas, R. S., Wesselkamper, S. C., Wang, N. C., Zhao, Q. J., Petersen, D. D., Lambert, J. C., Cote, I., Yang, L., Healy, E., Black, M. B., *et al.* (2013). Temporal concordance between apical and transcriptional points of departure for chemical risk assessment. *Toxicol. Sci*. **134,** 180–194.

Van den Berg, M., Birnbaum, L. S., Denison, M., De Vito, M., Farland, W., Feeley, M., Fiedler, H., Hakansson, H., Hanberg, A., Haws, L., *et al.* (2006). The 2005 World Health Organization reevaluation of human and Mammalian toxic equivalency factors for dioxins and dioxin-like compounds. *Toxicol. Sci*. **93,** 223–241.

Wambaugh, J. F., Setzer, R. W., Reif, D. M., Gangwal, S., Mitchell-Blackwood, J., Arnot, J. A., Joliet, O., Frame, A., Rabinowitz, J., Knudsen, T. B., *et al.* (2013). High-throughput models for exposure-based chemical prioritization in the ExpoCast project. *Environ. Sci. Technol*. **47,** 8479–8488.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Li, L., Clewell, H. J., 3rd, Judson, R. S., Freeman, K., Bao, W., Sochaski, M. A., Chu, T. M., *et al.* (2012a). Relative impact of incorporating pharmacokinetics on predicting in vivo hazard and mode of action from high-throughput in vitro toxicity assays. *Toxicol. Sci*. **132,** 327–346.

Wetmore, B. A., Wambaugh, J. F., Ferguson, S. S., Sochaski, M. A., Rotroff, D. M., Freeman, K., Clewell, H. J., 3rd, Dix, D. J., Andersen, M. E., Houck, K. A., *et al.* (2012b). Integration of dosimetry, exposure, and high-throughput screening data in chemical toxicity assessment. *Toxicol. Sci*. **125,** 157–174.